SUPPLEMENTARY INFORMATION:

MBNL proteins repress embryonic stem cell-specific alternative splicing and reprogramming

Hong Han[1,2] *, Manuel Irimia[1]*, P. Joel Ross[3], Hoon-Ki Sung[4], Babak Alipanahi[5], Laurent David[6], Azadeh Golipour[2,6], Mathieu Gabut[1], Iacovos P. Michael[4], Emil N. Nachman[1,2], Eric Wang[8], Dan Trcka[6], Tadeo Thompson[3], Dave O'Hanlon[1], Valentina Slobodeniuc[1], Nuno L. Barbosa-Morais[1,7], Christopher B. Burge[8], Jason Moffat[1,2], Brendan J. Frey[1,5], Andras Nagy[4,9], James Ellis[2,3], Jeffrey L. Wrana[2,6] and Benjamin J. Blencowe[1,2]

[1] Banting and Best Department of Medical Research and Donnelly Centre, University of Toronto, Toronto, Ontario, M5S 3E1, Canada
[2] Department of Molecular Genetics, University of Toronto, Toronto, Ontario, M5S 1A8, Canada
[3] Developmental and Stem Cell Biology, The Hospital for Sick Children, 101 College Street, Toronto, ON, M5G 1L7, Canada
[4] Center for Stem Cells and Tissue Engineering, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada
[5] Department of Electrical and Computer Engineering, University of Toronto, Canada
[6] Center for Systems Biology, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario, M5G 1X5, Canada
[7] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Portugal
[8] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[9] Department of Obstetrics and Gynecology, University of Toronto, Toronto, Ontario, M5S 1A8, Canada

* Co-first authors

Manuscript correspondence to:
Benjamin J. Blencowe, PhD
Donnelly Centre, University of Toronto
Toronto, ON, Canada
Office 416-978-3016; Fax 416-946-5545
Email: b.blencowe@utoronto.ca

SUPPLEMENTAL METHODS

List of siRNA target sequences (5'-3')

siRNAs targeting human MBNL1 (M-014136-01-0010)
GAAAATGATTGTCGGTTTG
ACAAGTATGTTACCCAGAT
TCACTGTGTGTATGGATTA
CCATAATATCTGCCGAACA

siRNAs targeting human MBNL2 (M-032198-00-0010)
GATGAAGAATGCAAATTTG
GCCACAAACTGCATGCTAA
GACAACACCGTAACCGTTT
GAGAGAGAACTGCAAGTAT

siRNAs targeting human MBNL3 (M-015373-00-0010)
CGTGATACCAAGTGGCTGA
GAGATTAATGGGCGGAACA
GGCTTTCAATCCTTACATA
GTAGAGAATTTCAGAGAGG

siRNAs targeting mouse Mbnl1 (M-065216-01-0010)
TAAATGGGCGGAATAACTT
GAAAGGTCGTTGCTCCAGA
CCACCAGGCTCAATATTGT
GACCAGACACGGAATGTAA
siRNAs targeting mouse Mbnl2 (M-065217-01-0010)
GGAACAATTTGATCCAGCA
CAGATGCAATTTATGTTTC
CTATTAGCTTTGCTCCTTA
ACAAACGACAACACCGTAA

Non-targeting siRNA pool (D-001810-10-20)

Mbnl1 and Mbnl2 overexpression in mouse ESCs
PiggyBac vectors "PB-TAG" and "PB-TAC" were generated to overexpress 3xFlag-Mbnl1 and 3xFlag-Mbnl2 proteins in mouse R1 ESCs. These vectors are a derivative of PB-TET [1] that co-express EGFP (PB-TAG) and mCherry (PB-TAC). Both vectors contain a tetracycline-regulated promoter controlling the expression of a bicistronic transgene comprising the gene of interest and EGFP or mCherry, respectively.

Overexpression R1 ESC lines were generated by co-transfection of: (1) PB-3xFlag-Mbnl1 and/or PB-3xFlag-Mbnl2, (2) PB-rtTA, a PB plasmid that includes the reverse tetracycline transactivator (rtTA) and the neomycin-resistance gene (neo)[2], and (3) a plasmid coding for piggyBac transposase (PBase)[3], followed by neomycin selection at 160μg/mL. A constitutive PB-EGFP, PB-rtTA and PBase co-transfection was used as a control.

The overexpression of 3xFlag-Mbnl1 and/or 3xFlag-Mbnl2 was induced by 1.5ug/mL of Dox, as determined by western blot analysis (Fig. S8a). To analyze the effect the 3xFlag-Mbnl1 and/or 3xFlag-Mbnl2 overexpression on ESC-differential AS, the cells were cultured with Dox for 48 hours. Cell populations with low and high EGFP and/or mCherry expression were collected on a MoFlo Astrios Flow Cytometer Cell sorter (Beckman Coulter) then lysed and analyzed by RT-PCR.

Mouse ESC differentiation

Multiple independent clonal mouse ESC (R1) line that co-express Mbnl1 and Mbnl2 at different levels under Dox inducible control were characterized. Following Dox induction, cells were cultured in mouse ESC media without LIF on 0.1% gelatin-coated plates. At day 6 following LIF withdrawal, expression of pluripotency factors in the clonal lines were analyzed by qRT-PCR and immunostaining. The parental R1 line, which was also treated with Dox, was used as a control.

Embryoid body (EB) differentiation was performed as described previously [4]. ESCs were trypsinized and plated in 10 cm non-adherent plates at a density of 5 X 10^5 cells per 10 mL in SFEB media. Cells were fed at days 3 and 5 by replacing 50% of the media. At day 7, the EB aggregates were seeded onto 0.1% gelatin-coated plates containing mouse ES media without LIF for another 4 days before harvesting for qRT-PCR analysis.

RNA-Seq data analysis and identification of ESC-differential AS events

Information on intron-exon structures was extracted from Ensembl annotations (release 65) for human (hg19) and mouse (mm9) genomes. In order to minimize the differences in annotation depth between mouse and human, mouse coordinates orthologous to human alternative exon-exon junctions (EEJs) were added to the annotations, using the lift-over tool from Galaxy (https://main.g2.bx.psu.edu/). From the resulting datasets, a Bowtie library of non-redundant EEJ sequences was generated for each species by combining every possible (forward combination) of splicing donor and acceptor within each gene. For each EEJ sequence, we determined the effective number of unique mappable positions. We extracted the L-k+1 (L being the EEJ length) k-mers from each EEJ sequence and then aligned the full set of k-mers against the EEJ library plus the respective genome using Bowtie [5], allowing for a maximum of two mismatches along the entire length of the read. The number of k-mers with one unique alignment was counted for each EEJ; this corresponds to the junction's effective number of unique mappable positions for a given set of RNA-Seq k-mers.

RNA-Seq reads from the different samples were then mapped to the EEJ libraries using Bowtie [5] with –m 1 –v 2 parameters. Reads were trimmed to 50 nucleotides, if longer, and reads mapping to the genome were previously discarded (since bona fide EEJs do not exist in the genome). A minimum of eight mapped nucleotides overlap was required across a given EEJ. Next, the outputs were parsed to

identify cassette alternative exons (i.e. exons that are either included or fully excluded from the transcripts) by identifying exons that have associated reads mapping to (i) both EEJs formed by inclusion of the exon (i.e. a constitutive upstream exon [C1] joined to an alternative exon [A] and an [A] exon joined to a constitutive downstream exon [C2]), referred to as C1A and AC2 junctions below, and (ii) the EEJ formed by exclusion of the exon (C1C2).

The inclusion level of a cassette alternative exon was defined as the *percentage* of gene transcripts in which the exon is *spliced in* (PSI). To calculate this, we used a two-step approach. First, for each possible combination of C1 and C2 exons, PSI was estimated using read counts mapping to EEJs, corrected for the mappable positions ($EEJ_{count} * \frac{Maximum_{mappability}}{EEJ_{mappability}}$), following the formula:

$$PSI = 100 * \frac{\mathrm{C1A} + \mathrm{AC2} + \Sigma(\mathrm{C}_i\mathrm{A}) + \Sigma(\mathrm{AC}_j)}{\mathrm{C1A} + \mathrm{AC2} + \Sigma(\mathrm{C}_i\mathrm{A}) + \Sigma(\mathrm{AC}_j) + 2 * (\mathrm{C1C2} + \Sigma(\mathrm{C}_i\mathrm{C2}) + \Sigma(\mathrm{C1C}_j))}$$

where C_i is any possible splice donor upstream of the alternative exon (not including C1), and C_j any possible splice acceptor downstream of the alternative exon (not including C2). A reference pair of C1 and C2 exons was then selected (if more than one) as the combination yielding the maximum number of reads for the sum of the C1A, AC2 and C1C2 junctions. Alternative splicing events were only included when they met a minimal transcript coverage requirement of (i) ≥15 corrected reads mapping to the sum of exclusion EEJs, or (ii) ≥15 mappability-corrected reads mapping to one of the two inclusion EEJs ($C1A+C_iA$ or AC_j+AC2), and ≥10 to the other inclusion EEJ. For alternative exons with multiple acceptor/donor splice sites, we used the splice site combination with the highest read support. A similar approach to calculate PSIs was recently used by Pervouchine *et al.* [6].

REFERENCES

1. Woltjen, K. *et al.* piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* **458**, 766-70 (2009).
2. Li, Z., Michael, I.P., Zhou, D., Nagy, A. & Rini, J. A simple piggyBac transposon-based mammalian cell expression system for inducible protein production. *Proc. Natl. Acad. Sci. U.S.A.* **In press**(2013).
3. Wang, W. *et al.* Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 9290-5 (2008).
4. Watanabe, K. *et al.* Directed differentiation of telencephalic precursors from embryonic stem cells. *Nature neuroscience* **8**, 288-96 (2005).
5. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**(2009).
6. Pervouchine, D.D., Knowles, D.G. & Guigo, R. Intron-centric estimation of alternative splicing from RNA-seq data. *Bioinformatics* **29**, 273-4 (2013).

TABLE LEGENDS

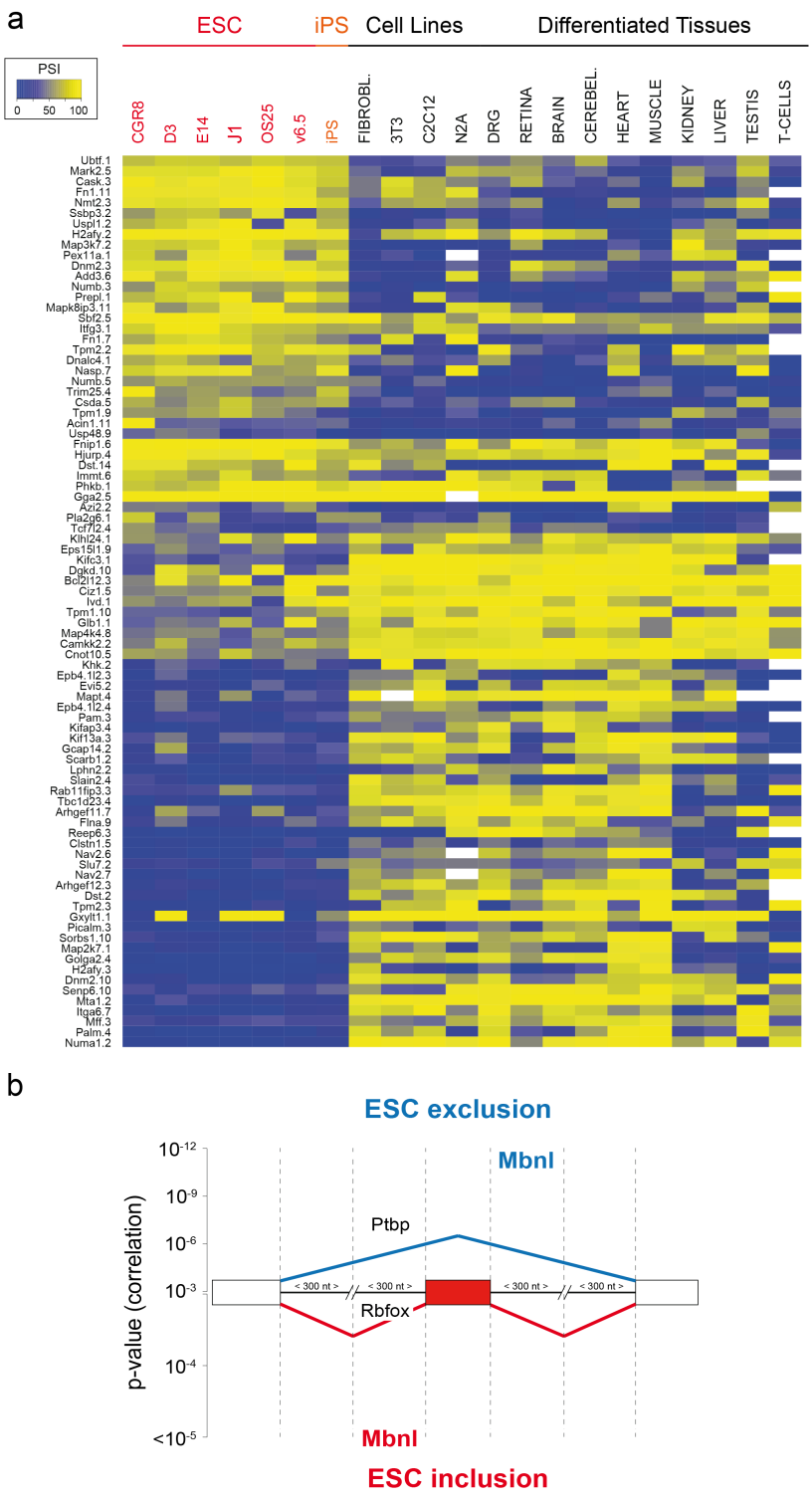
Supplementary Table 1 – Information on RNA-Seq datasets and samples
Read length, length used in the bowtie alignments; # Reads, number of unique reads or read pairs in each sample; Source (SRA), identifiers for the NCBI Short Read Archive, where available; Reference/GEO, PubMed ID (where published available) or Gene Expression Omnibus (GEO) accession.

Supplementary Table 2 – Information on Human and mouse ESC-differential AS events
EX LENGTH, exon length; FullCO, full coordinates of the AS event (chromosome, C1 donor, AS exon, C2 acceptor); av/min/max ESC/DIF/CL/iPS, average/minimum/maximum PSI in ESCs/differentiated tissues/cell lines/iPS cells; difESC_DIF/CL, PSI difference between ESC and differentiated tissues/cell lines; Dif_KD_HeLa/293T/C2C12, PSI difference between siMBNL1+2 and siControl in HeLa/293T/C2C12 cells; C1/A/C2 sequence of C1, A and C2 exons of each event.

Supplementary Table 3 – DAVID (david.abcc.ncifcrf.gov/) output for functional enrichment categories for human, mouse or conserved ESC-differential AS events
Three major functional clusters that were consistently enriched in genes harbouring the three groups of AS events are indicated: cytoskeleton (green), plasma membrane (yellow), associated to protein kinase (orange).

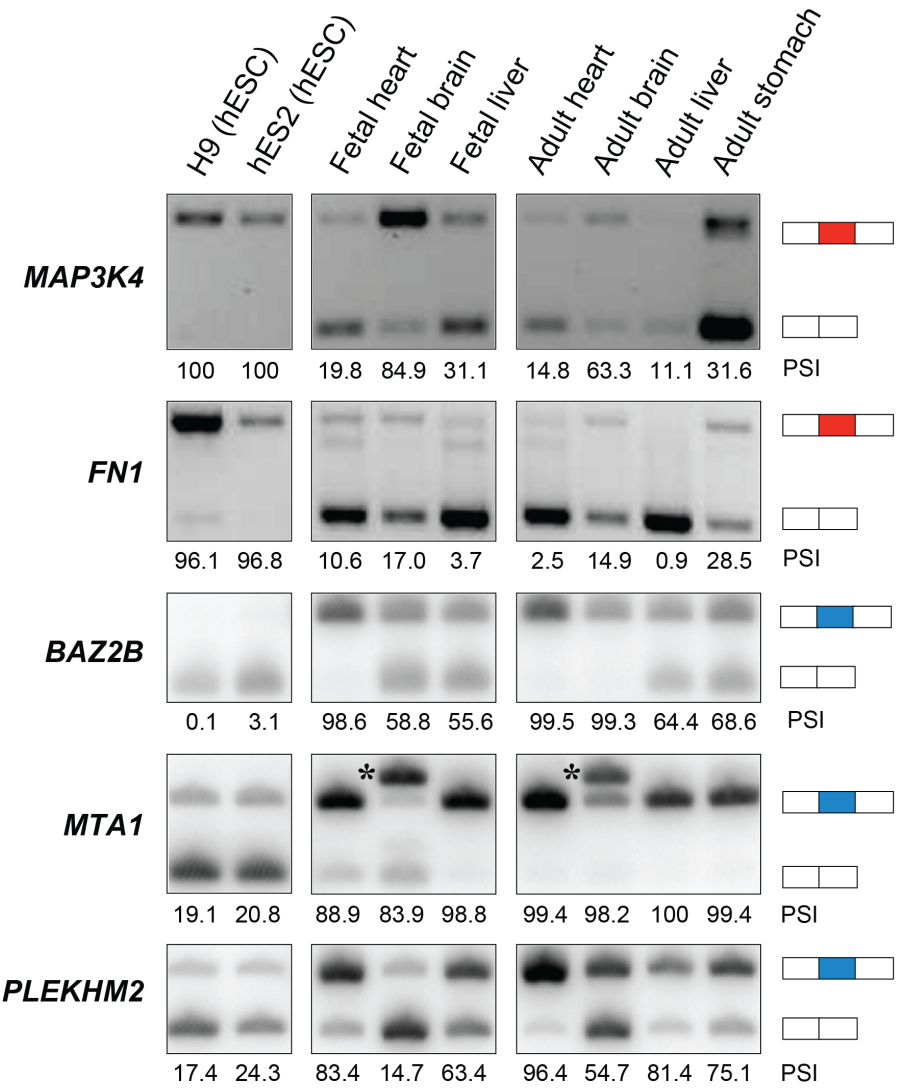
Supplementary Table 4 – Expression levels of the human and mouse splicing factors analyzed by RNA-Seq
ESC-rank, rank of differential expression between ESC/iPSCs and differentiated tissues and cell lines (CL); Av_ESC+iPSCs, cRPKM average for the ESC and iPSC samples (columns H-N); Av_DIF, cRPKM average for the differentiated tissues and cell lines samples (columns M onwards); C_p-value Wilcox, Bonferroni-corrected p-value of Wilcox test of ESC/iPSCs versus differentiated tissues and cell lines after quantile normalization.

Supplementary Table 5 – Information on mouse ESC-differential AS events plotted in Supplementary Fig. 14
AS events shown were represented by sufficient read coverage and ≥ 25 PSI differences between iPSCs and MEFs, as described above. TD/TI: refers to whether the clones died after Dox removal (transgene-dependent, n=5) or become stable iPSCs (transgene-independent, n=3); 21D refers to the day following Dox induction.



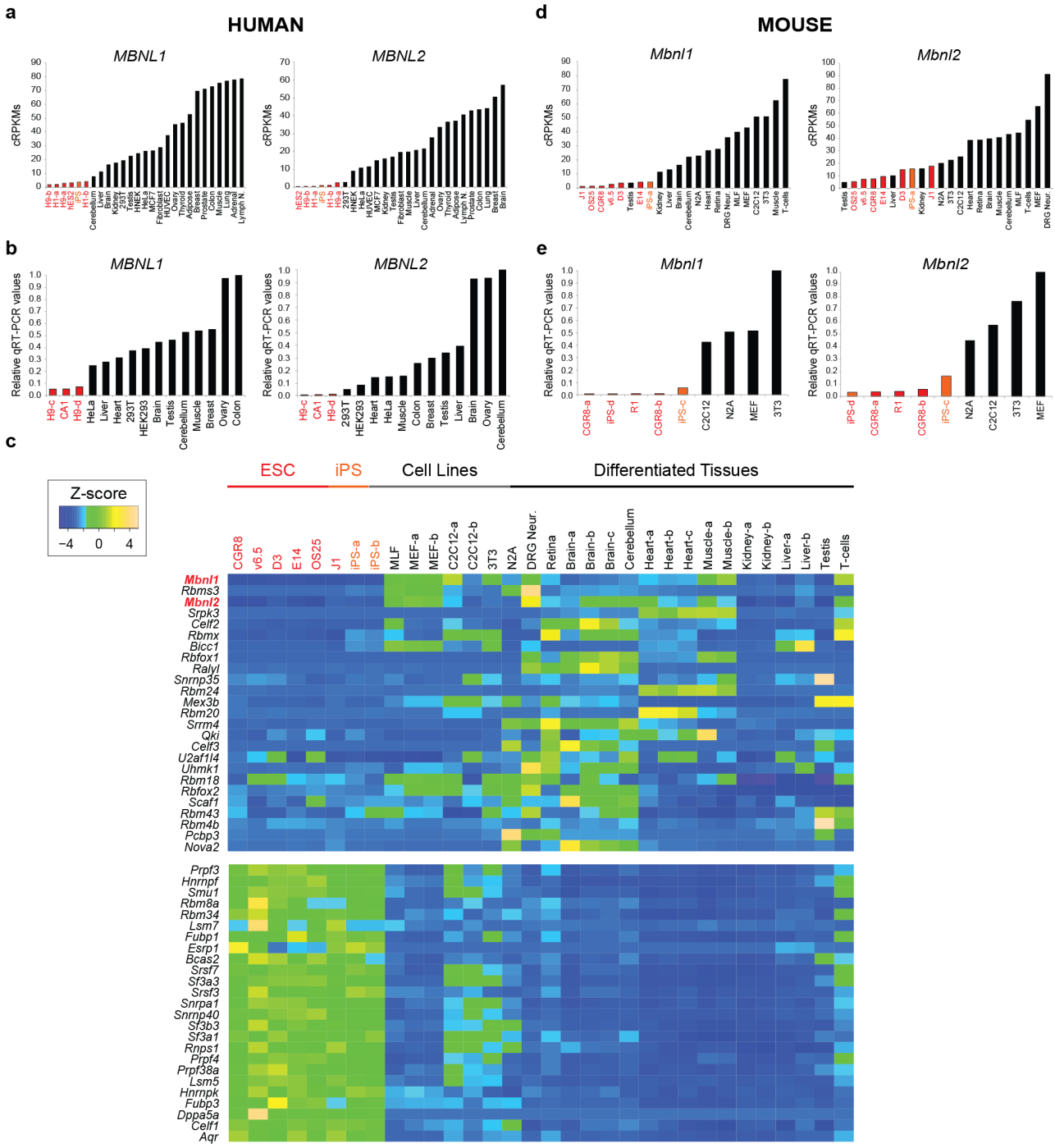
Supplementary Fig. 1 – Identification of ESC-differential AS (mouse)

a) Heatmap of percent spliced in (PSI) values for 85 representative mouse ESC-differential AS events in transcripts that are widely expressed across ESCs, iPSCs, non-ESC lines and differentiated tissues. b) Splicing code features that are significantly associated ($p<0.001$, Pearson correlation) with mouse ESC-differential AS. Features are ranked according to Pearson correlation p-values (y-axis) for alternative exons with either lower (top) or higher (bottom) inclusion in ESCs/iPSCs.



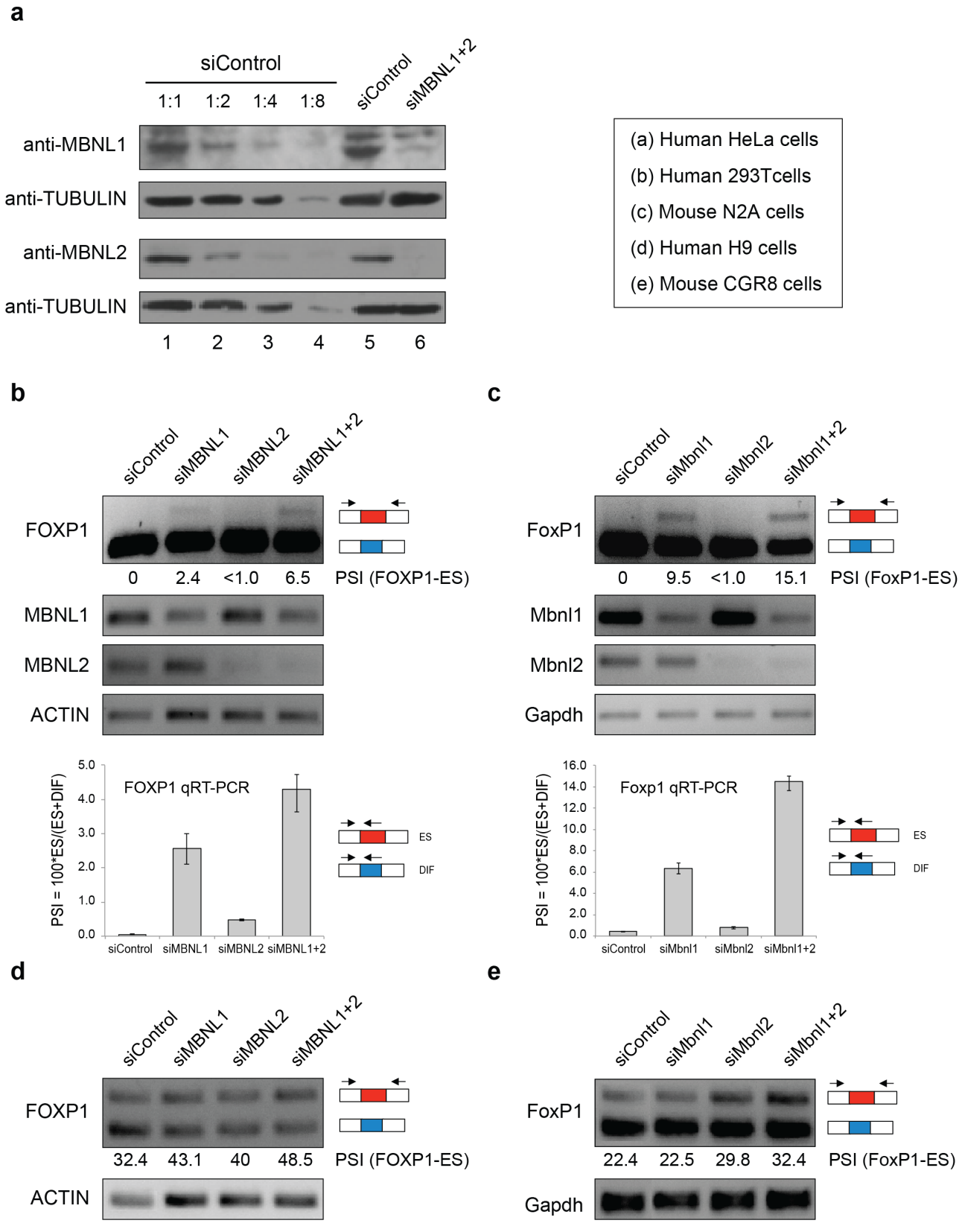
Supplementary Fig. 2 – Examples of human ESC-differential AS events

Representative RT-PCR assays monitoring AS patterns of *MAP3K4, FN1, BAZ2B, MTA1* and *PLEKHM2* transcripts in human ESCs (H9 and hES2), fetal (heart, brain and liver) and adult (heart, brain, liver and stomach) tissues. Red and blue exons indicate those exons with higher and lower inclusion in ESCs compared to differentiated tissues/cell lines, respectively. * indicates brain-specific MTA1 isoform.



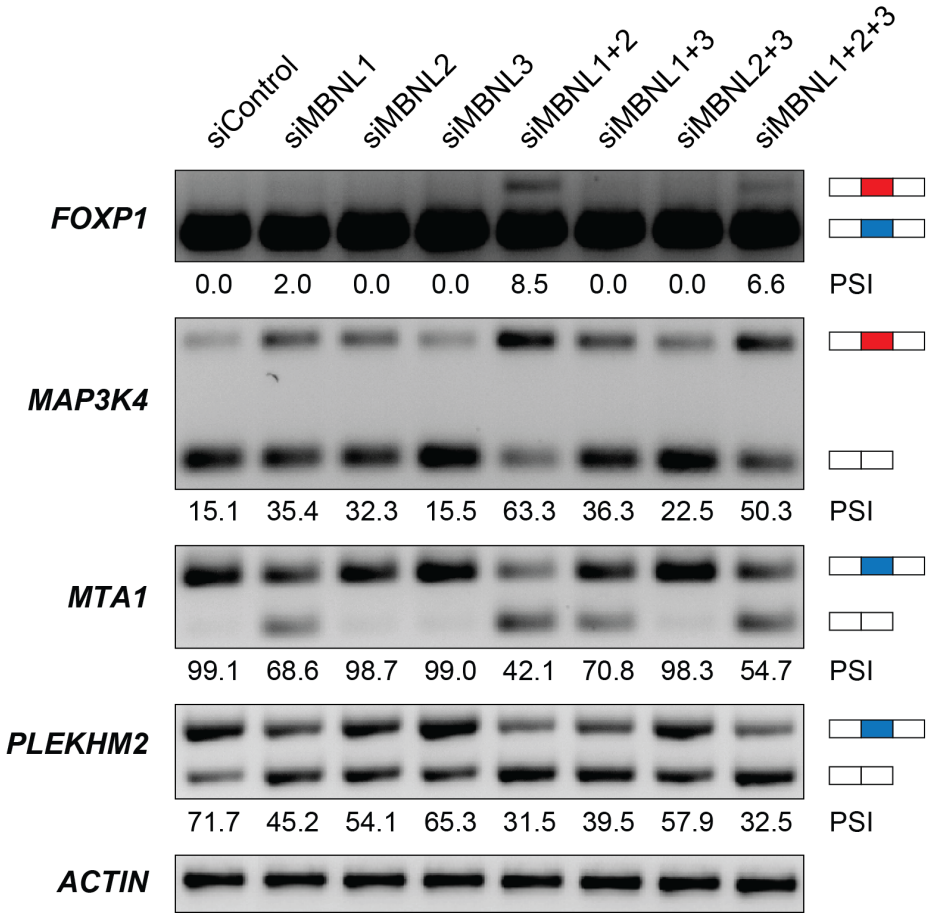
Supplementary Fig. 3 – mRNA expression profiling of human and mouse splicing factors

a,b) RNA-Seq and qRT-PCR data showing expression levels of human *MBNL1* and *MBNL2* mRNAs. c) Heatmap of Z-scores of mRNA expression levels for the 25 mouse splicing factors with the lowest or highest overall rank of ESC mRNA expression compared to differentiated tissues and cell lines. Additional information in Supplementary Table 4. d,e) RNA-Seq and qRT-PCR data showing expression levels of mouse *Mbnl1* and *Mbnl2* mRNAs.


a
siControl
1:1 1:2 1:4 1:8
siControl
siMBNL1+2
anti-MBNL1
anti-TUBULIN
anti-MBNL2
anti-TUBULIN
1 2 3 4 5 6
(a) Human HeLa cells
(b) Human 293Tcells
(c) Mouse N2A cells
(d) Human H9 cells
(e) Mouse CGR8 cells
b
siControl
siMBNL1
siMBNL2
siMBNL1+2
FOXP1
0 2.4 <1.0 6.5
PSI (FOXP1-ES)
MBNL1
MBNL2
ACTIN
FOXP1 qRT-PCR
PSI = 100*ES/(ES+DIF)
5.0
4.0
3.0
2.0
1.0
0.0
siControl siMBNL1 siMBNL2 siMBNL1+2
ES
DIF
c
siControl
siMbnl1
siMbnl2
siMbnl1+2
FoxP1
0 9.5 <1.0 15.1
PSI (FoxP1-ES)
Mbnl1
Mbnl2
Gapdh
Foxp1 qRT-PCR
PSI = 100*ES/(ES+DIF)
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
siControl siMbnl1 siMbnl2 siMbnl1+2
ES
DIF
d
siControl
siMBNL1
siMBNL2
siMBNL1+2
FOXP1
32.4 43.1 40 48.5
PSI (FOXP1-ES)
ACTIN
e
siControl
siMbnl1
siMbnl2
siMbnl1+2
FoxP1
22.4 22.5 29.8 32.4
PSI (FoxP1-ES)
Gapdh

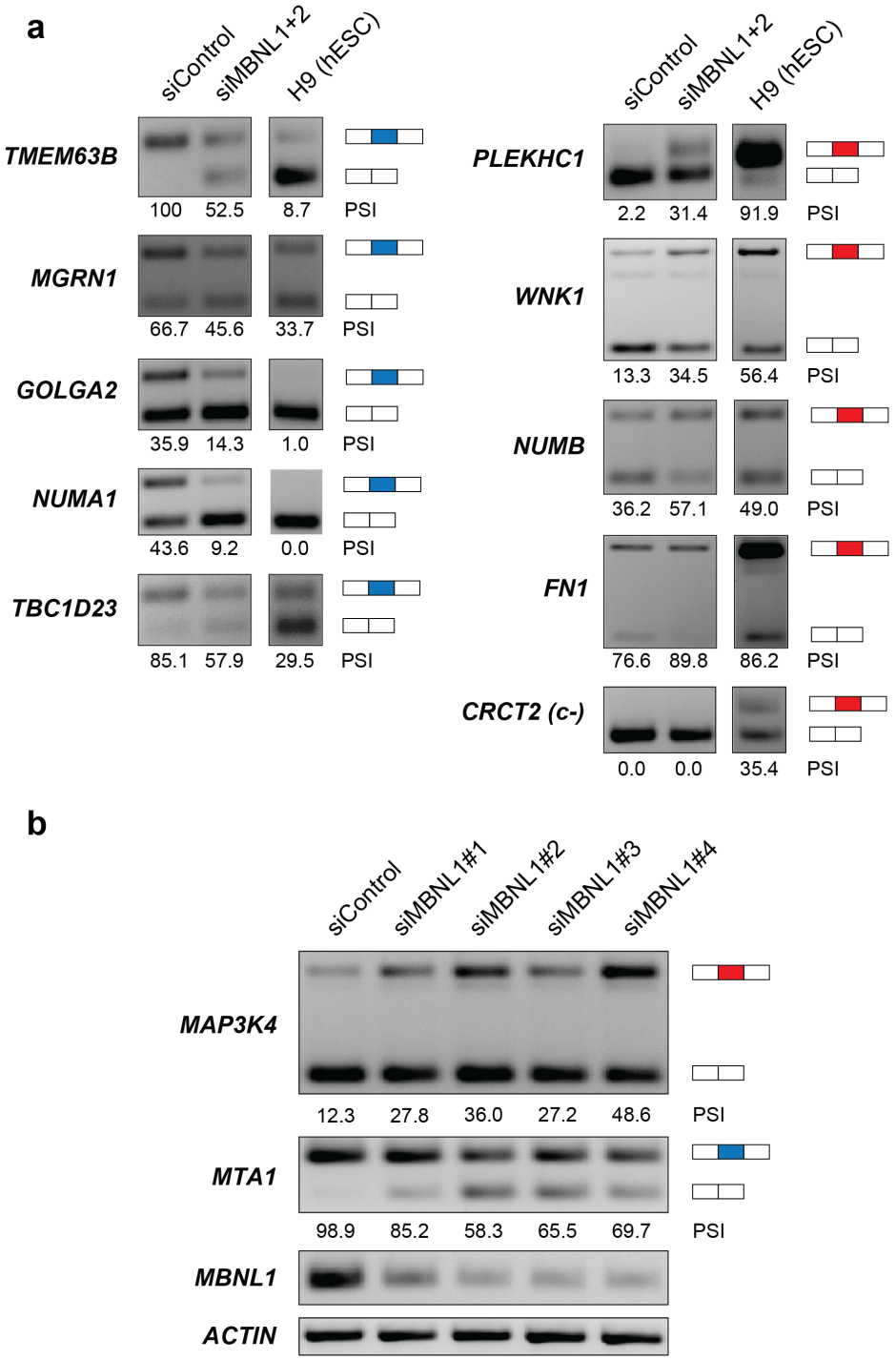
Supplementary Fig. 4 – siRNA knockdown of MBNL1 and MBNL2 promotes ESC-specific AS of FOXP1 transcripts

a) Western blot confirming efficient knockdown of MBNL1 and MBNL2 proteins in HeLa cells transfected with siRNA pools targeting these factors (siMBNL1+2, lane 6). Lane 5, lysate from cells transfected with a non-targeting siRNA pool (siControl). Lanes 1-4, serial dilutions (1:1, 1:2, 1:4 and 1:8) of lysate from cells transfected with siControl. b) RT-PCR assays monitoring mRNA levels and/or AS patterns of FOXP1, MBNL1 and MBNL2 transcripts in 293T cells transfected with control siRNAs (siControl) or siRNA pools targeting MBNL1 (siMBNL1), MBNL2 (siMBNL2) or both of these factors (siMBNL1+2). Primers specific to the constitutive exons were used to amplify both FOXP1 isoforms. "PSI values" based on the ratios of expression of FOXP1 canonical (blue exon) and FOXP1-ES (red exon) isoforms analyzed by qRT-PCR using splice junction-specific primers are shown below. c) mRNA levels and/or AS patterns of murine Foxp1, Mbnl1 and Mbnl2 were assayed as in (b) following transfections of control siRNAs (siControl) or siRNAs targeting Mbnl1 (siMbnl1), Mbnl2 (siMbnl2) or both of these factors (siMbnl1+2) in mouse N2A cells. d,e) mRNA levels and/or AS patterns of FOXP1/Foxp1 were assayed as in (b and c) following transfections of siRNA pools in human H9 ESCs (d) mouse CGR8 ESCs (e) respectively.



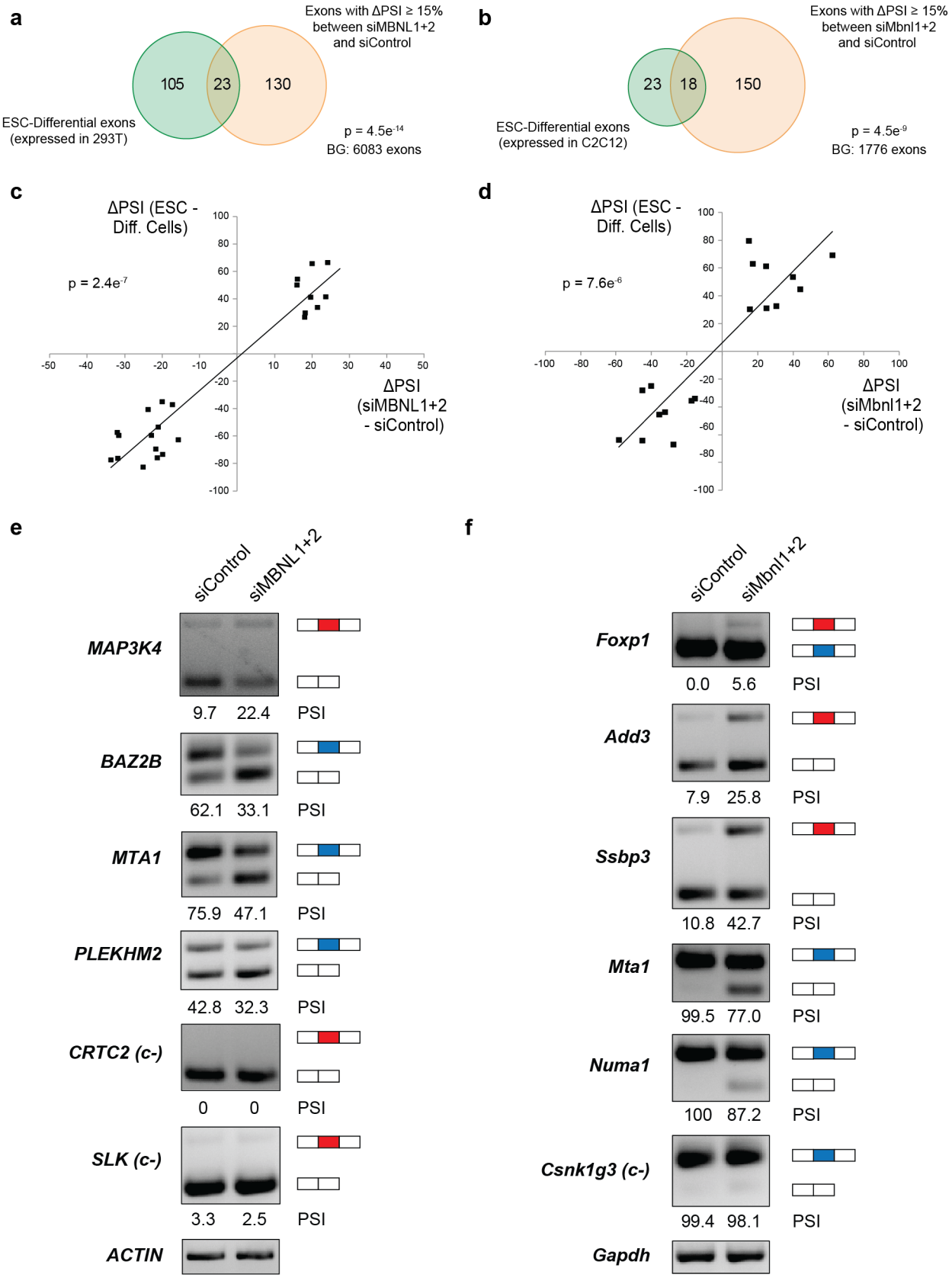
Supplementary Fig. 5 – Comparison of effects of siRNA knockdown of MBNL1, MBNL2 and MBNL3 on ESC-differential AS

Representative ESC-differential AS events were analyzed by RT-PCR assays following transfection of HeLa cells with a control siRNA, or with single or combinations of siRNA pools that target MBNL1, MBNL2 and MBNL3 (refer also to Supplementary Fig. 4 legend).



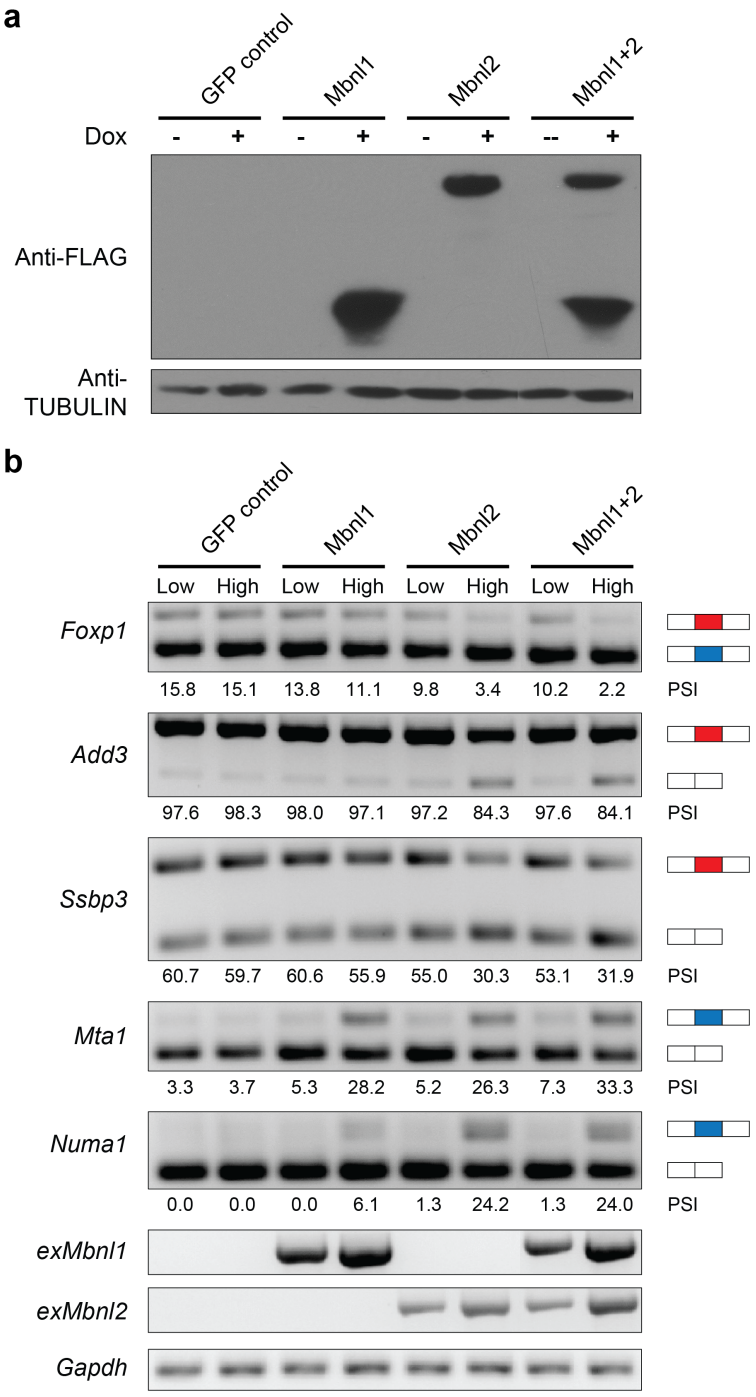
Supplementary Fig. 6 – Additional examples of ESC-differential AS events regulated by MBNL1 and MBNL2

a) Representative RT-PCR validation assays for ESC-differential AS events that have PSI changes in HeLa cells following siMBNL1+2 transfection (refer to Supplementary Fig. 4 legend); splicing patterns in human H9 ESCs are shown for comparison. b) RT-PCR analysis of ESC-differential AS events following transfection of single siRNAs comprising the siRNA pool targeting MBNL1 used in Figs. 3, S5 and S6. Results for representative ESC-differential AS events are shown (compare also with results in Supplementary Fig. 5).


a
Exons with ΔPSI ≥ 15% between siMBNL1+2 and siControl
105
23
130
ESC-Differential exons (expressed in 293T)
p = 4.5e-14
BG: 6083 exons
b
Exons with ΔPSI ≥ 15% between siMbnl1+2 and siControl
23
18
150
ESC-Differential exons (expressed in C2C12)
p = 4.5e-9
BG: 1776 exons
c
ΔPSI (ESC - Diff. Cells)
p = 2.4e-7
ΔPSI (siMBNL1+2 - siControl)
d
ΔPSI (ESC - Diff. Cells)
p = 7.6e-6
ΔPSI (siMbnl1+2 - siControl)
e
siControl
siMBNL1+2
MAP3K4
9.7 22.4 PSI
BAZ2B
62.1 33.1 PSI
MTA1
75.9 47.1 PSI
PLEKHM2
42.8 32.3 PSI
CRTC2 (c-)
0 0 PSI
SLK (c-)
3.3 2.5 PSI
ACTIN
f
siControl
siMbnl1+2
Foxp1
0.0 5.6 PSI
Add3
7.9 25.8 PSI
Ssbp3
10.8 42.7 PSI
Mta1
99.5 77.0 PSI
Numa1
100 87.2 PSI
Csnk1g3 (c-)
99.4 98.1 PSI
Gapdh

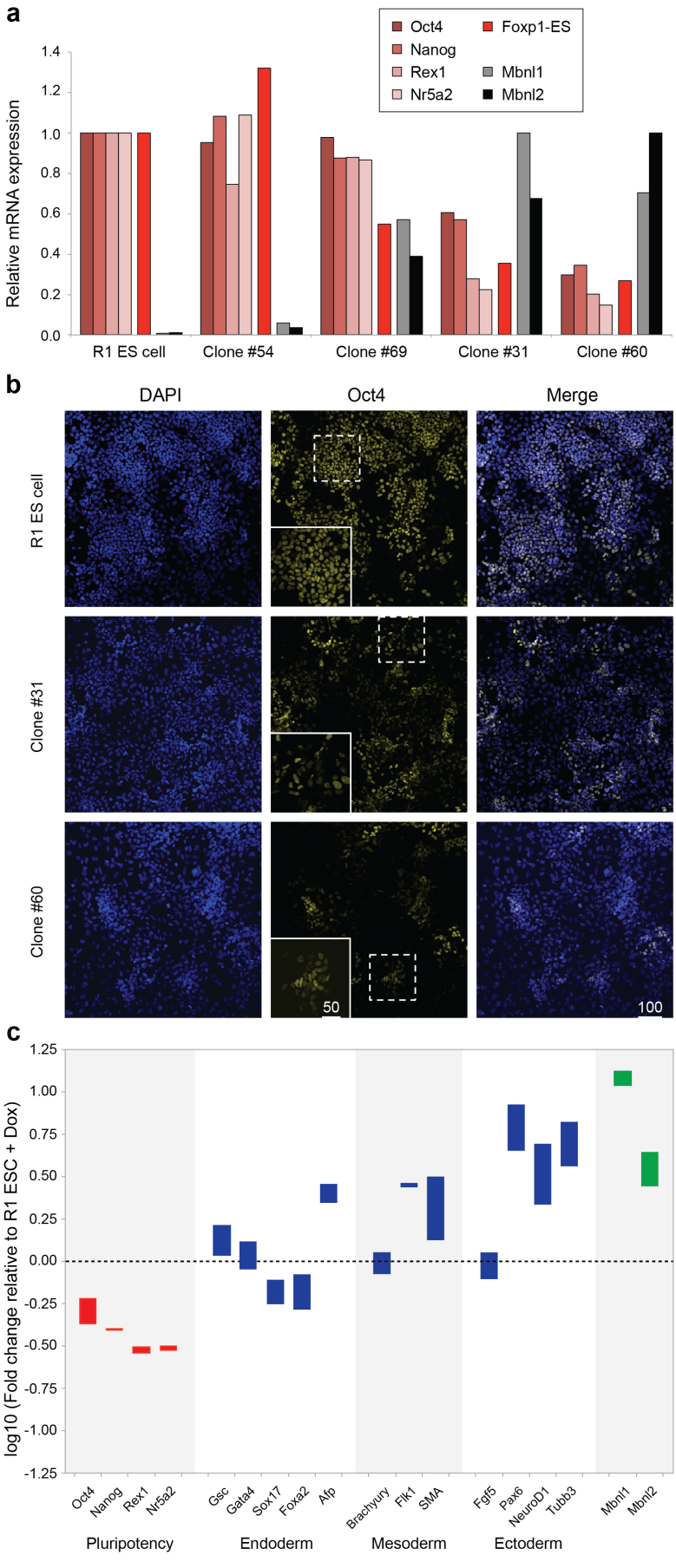
Supplementary Fig. 7 – Analysis of ESC-differential AS events affected by knockdown of MBNL1 and MBNL2 in human 293T and mouse C2C12 cells

a,b) Venn diagrams showing proportions of human and mouse ESC-differential AS events (green) that display ≥15 PSI change between 293T or C2C12 cells transfected with siMbnl1+2 versus a siControl control (refer to Supplementary Fig. 4 legend) (orange). c,d) High association ($p < 2.4e^{-7}$ (c) and $p < 7.6e^{-6}$ (d), one-sided binomial tests between quadrants) between differences in PSIs of ESCs versus differentiated cells/tissues, and between differences in PSIs of siMbnl1+2 knockdown versus siControl treatments of human 293T (c) or mouse C2C12 cells (d). Linear regression lines are shown. e,f) RT-PCR validation of RNA-Seq-detected changes in PSI for ESC-differential AS events, following siMbnl1+2 knockdown in 293T (e) and in C2C12 cells (f). ESC-differential AS events detected by RNA-Seq not to change upon siMbnl1+2 knockdown are included as specificity controls (c-): CRTC2, SLK in (e) and Csnk1g3 in (f).



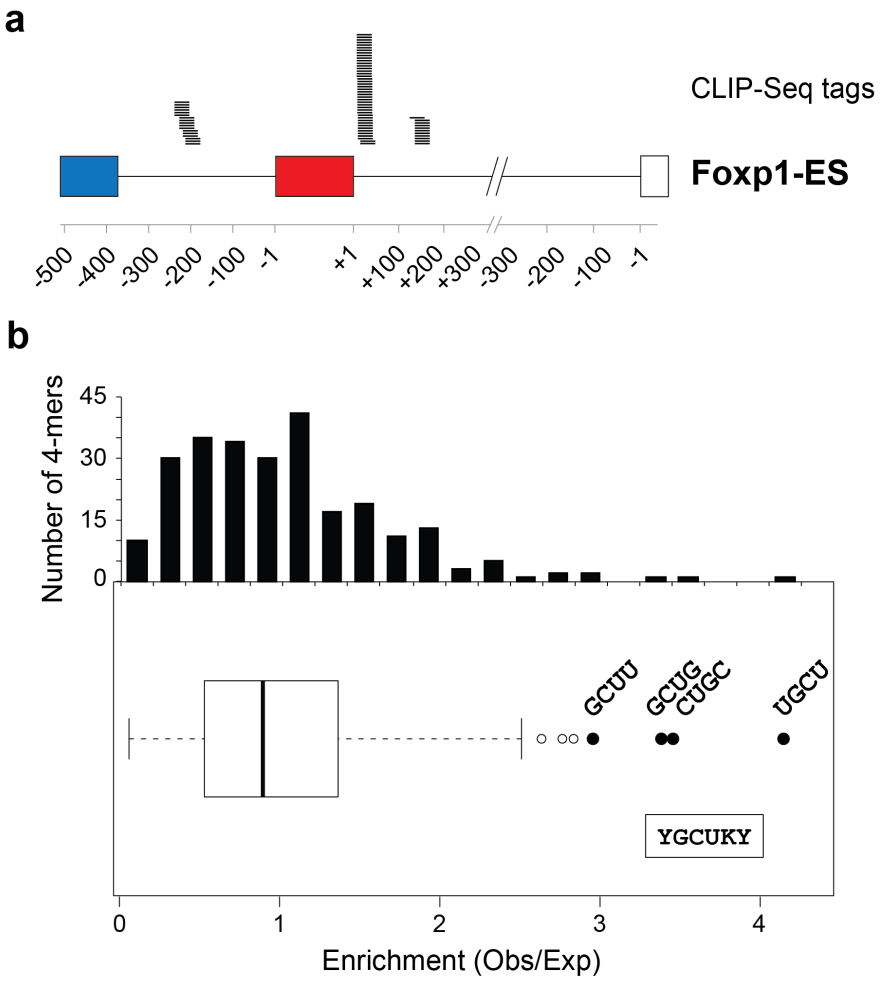
Supplementary Fig. 8 – Overexpression of Mbnl1 and Mbnl2 in ESCs promotes differentiated cell-like AS patterns for ESC-differential AS events

a) Cell lysates prepared from mouse R1 ESCs over-expressing 3xFlag-Mbnl1 and/or 3xFlag-Mbnl2 under Dox inducible control, or EGFP ("GFP control") were immunoblotted and probed with anti-Flag antibody (refer to Methods for details). b) RNA isolated from cell populations with low and high EGFP and/or mCherry expression (collected by flow cytometer cell sorter) was analyzed by RT-PCR to assess effects of Mbnl proteins on ESC-differential AS events. Exogenous (ex) expression of Mbnl1 and Mbn2 transcripts was analyzed by RT-PCR in the same samples, as indicated.


a
Oct4
Nanog
Rex1
Nr5a2
Foxp1-ES
Mbnl1
Mbnl2
Relative mRNA expression
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
R1 ES cell
Clone #54
Clone #69
Clone #31
Clone #60
b
DAPI
Oct4
Merge
R1 ES cell
Clone #31
Clone #60
50
100
c
log10 (Fold change relative to R1 ESC + Dox)
1.25
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
-1.25
Oct4
Nanog
Rex1
Nr5a2
Gsc
Gata4
Sox17
Foxa2
Afp
Brachyury
Flk1
SMA
Fgf5
Pax6
NeuroD1
Tubb3
Mbnl1
Mbnl2
Pluripotency
Endoderm
Mesoderm
Ectoderm

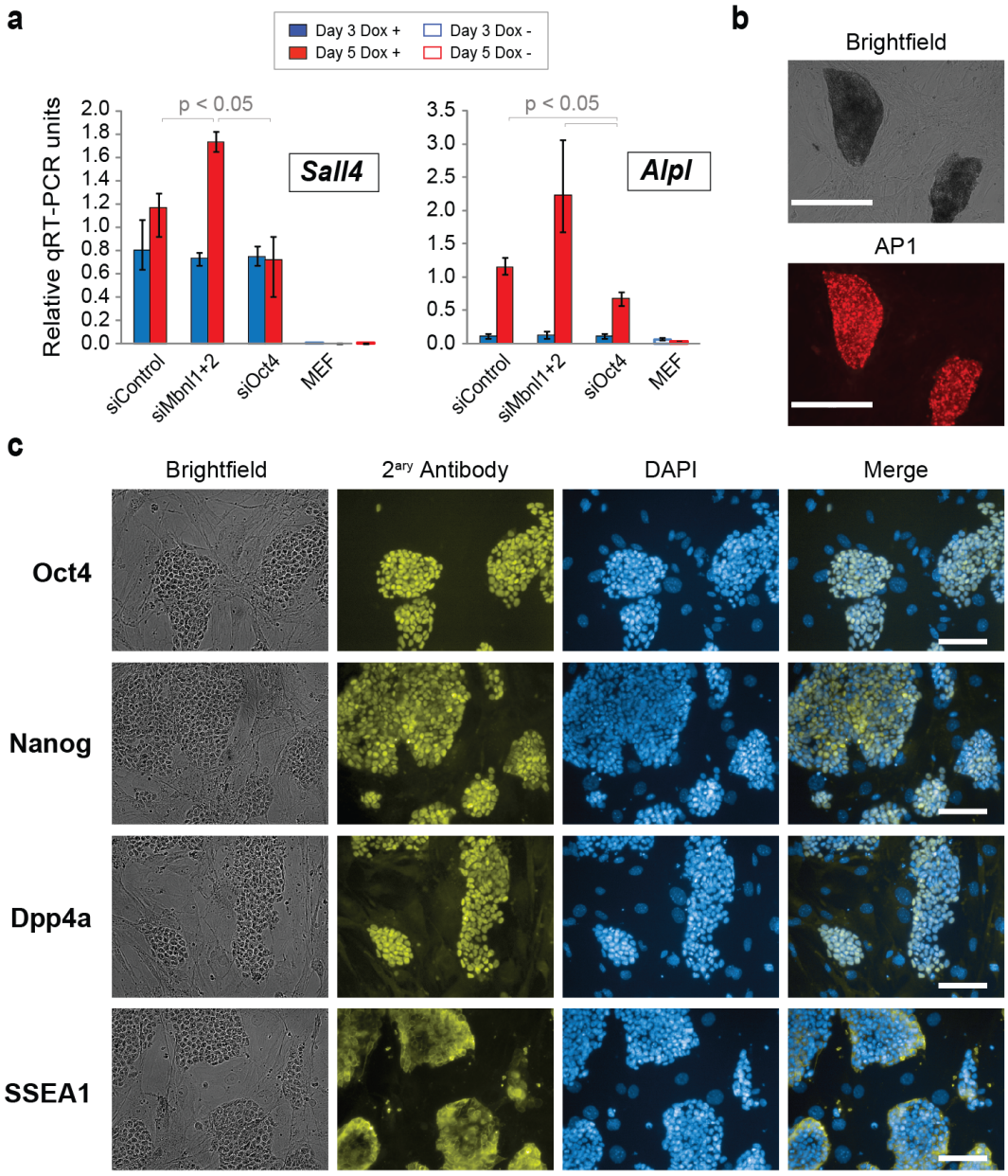
Supplementary Fig. 9 – Overexpression of Mbnl proteins suppresses pluripotency factors and promotes expression of differentiation markers

a) Expression of pluripotency factors and Foxp1-ES isoform (red bars) was analyzed by qRT-PCR in independent clonal mouse ESC (R1) lines overexpressing Mbnl1+2 and the parental R1 line at Day 6 of LIF withdrawal. The total levels of Mbnl1 and Mbnl2 are indicated as grey and black bars respectively. b) Representative images of the parental R1 and independent overexpression lines at Day 6 of LIF withdrawal immunostained with antibody specific for Oct4 and counter stained with DAPI. c) Expression of pluripotency factors (Oct4, Nanog, Rex1 and Nr5a2) and differentiation markers specific for endoderm (Gsc, Gata4, Sox17, Foxa2 and Afp), mesoderm (Brachyury, Flk1 and SMA), and ectoderm (Fgf5, Pax6, NeuroD1 and Tubb3) was analyzed by qRT-PCR at day 11 of EB differentiation. Red, blue and green bars show the range of expression levels for pluripotency, differentiation markers, and total Mbnl1/2, in independent clonal overexpression lines respectively, relative to the parental R1 line treated with Dox (normalized to 1.0).



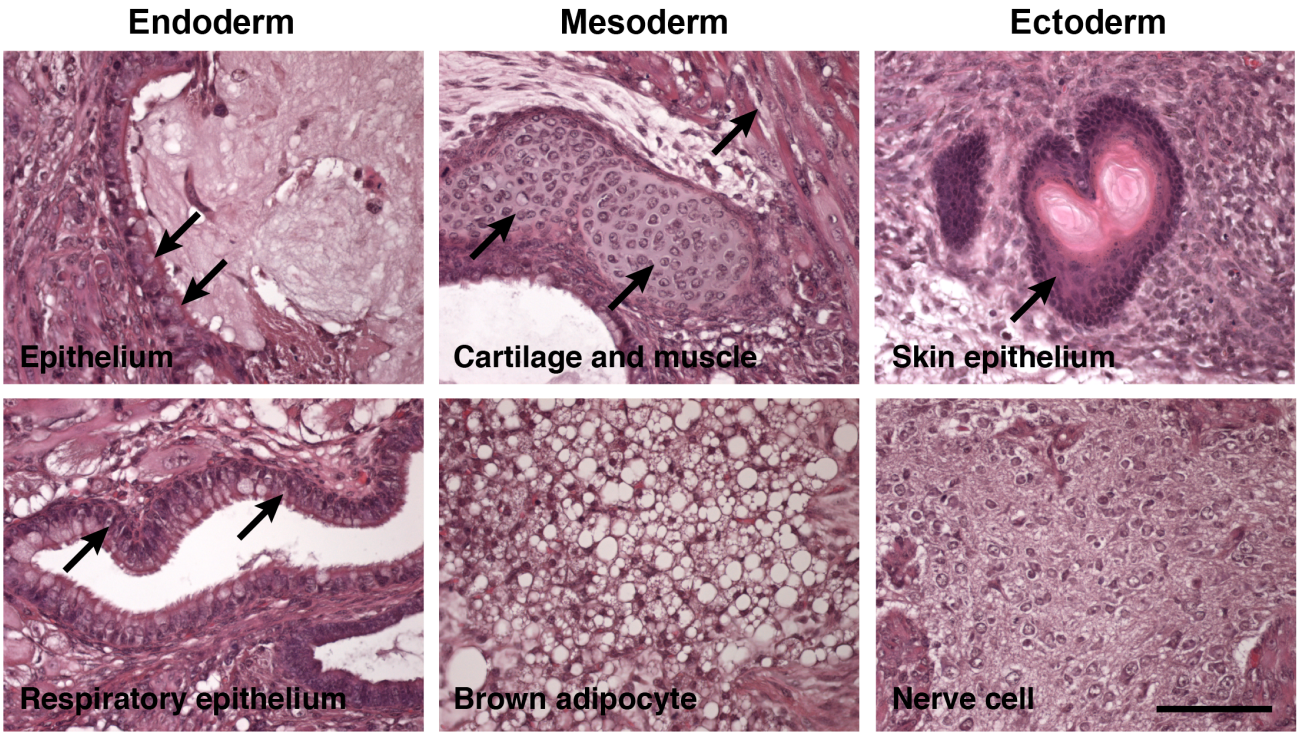
Supplementary Fig. 10 – Mapping of Mbnl1 CLIP-Seq tags in the Foxp1 gene and sequence motifs enriched under Mbnl1 CLIP-Seq clusters

a) CLIP-Seq tags for Mbnl1 from murine myoblast C2C12 cells are shown mapped to the vicinity of Foxp1 exon 16b, the inclusion of which forms the Foxp1-ES isoform. The highest density of mapped tags is adjacent to the 5′ splice site of exon 16b, and binding of Mbnl1 in this region in differentiated cells is predicted to interfere with recognition by U1 snRNP. b) Histogram and box-plot of the ratios of observed versus expected occurrences of 4-mers in Mbnl1 CLIP clusters in ESC-differential events in C2C12 cells, relative to randomized sequences from the same regions. An Mbnl RNA binding consensus (boxed) derived from the literature (refer to main text) is shown for comparison.



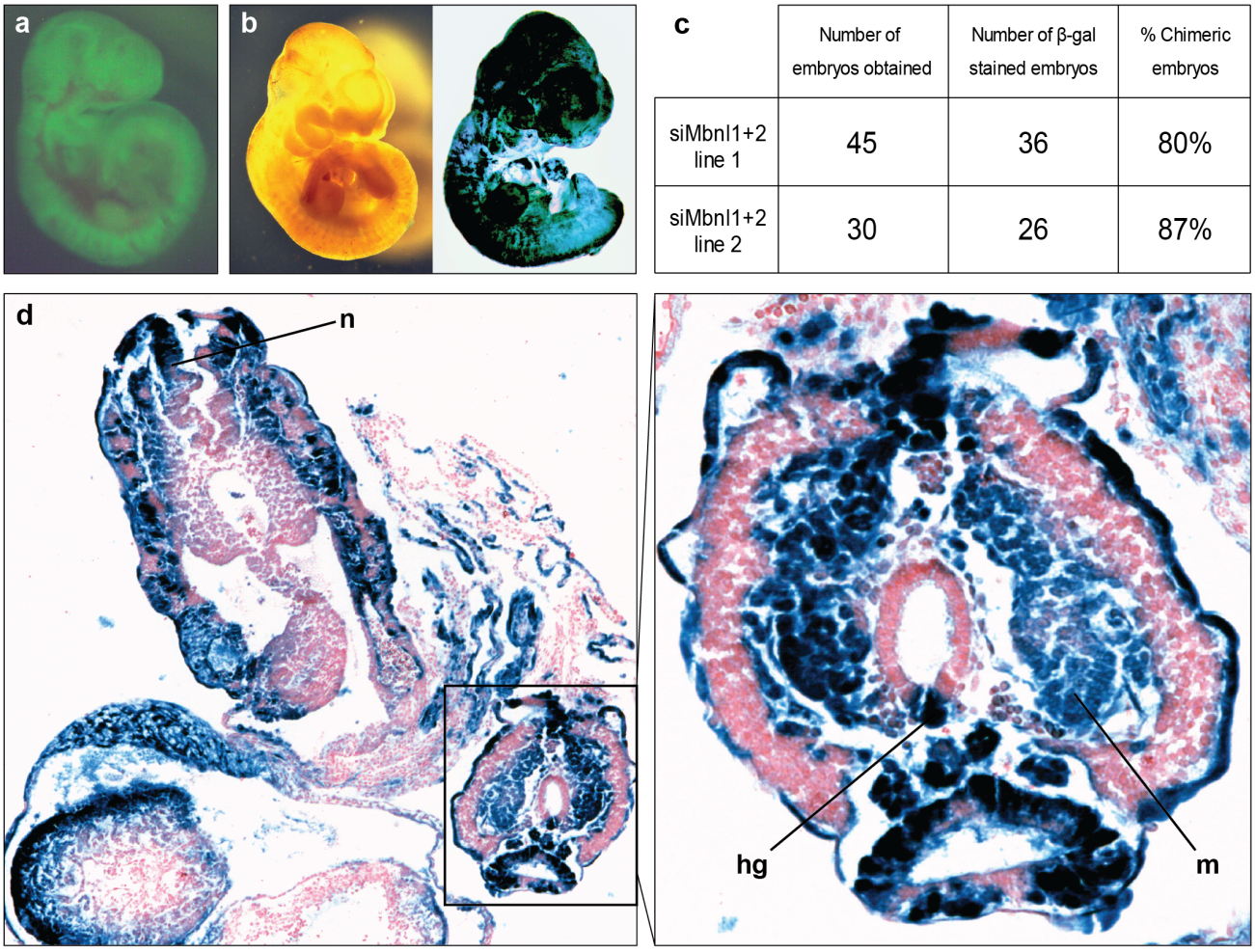
Supplementary Fig. 11 – Characterization of secondary MEF reprogramming following siRNA knockdown of Mbnl proteins

a) qRT-PCR quantification of mRNA expression levels of endogenous Sall4 and Alpl genes during reprogramming. Secondary MEFs were transfected with control siRNAs (siControl) or siRNAs targeting Mbnl1 and Mbnl2 (siMbnl1+2) or Oct4 (siOct4) and treated with doxcycline (Dox) for 3 days (blue bars) or 5 days (red bars) before analysis. Secondary MEFs without Dox induction are shown as empty bars. Values represent Means ± Range (n=3). t-test was used to asses statistical significance of the differences between conditions. b) Secondary MEFs were transfected with Mbnl1 and Mbnl2 siRNAs (siMbnl1+2) and treated with Dox for 8 days followed by 5 days of Dox withdraw. The resulting iPSC colonies were stained with alkaline phosphatase (AP) at day 13 and representative microscope images are shown (10X, scale bar = 400 μm). c) Representative images of Dox-independent iPSC colonies immunostanined with antibodies specific for Oct4, Nanog, Dpp4a and SSEA1 and counter stained with DAPI. Scale bar = 100 μm.



Supplementary Fig. 12 – Teratoma analysis assessing the pluripotency potential of siMbnl1+2 iPSC lines derived from transgene-independent colonies during secondary MEF reprogramming

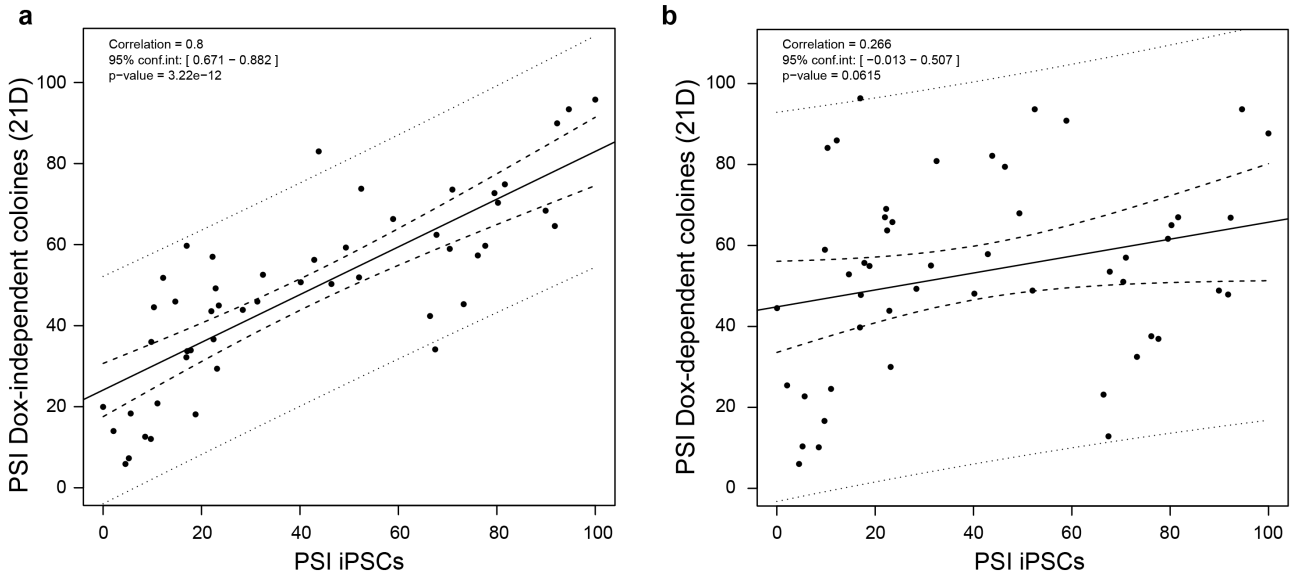
Representative sections from teratomas generated from two independent transgene-independent siMbnl1+2 iPSC lines generated by secondary MEF reprogramming. Sections were stained with hematoxylin and eosin. Cell types representing all three germ layers – endoderm, mesoderm, and ectoderm – are indicated. Scale bar = 100 μm.



	Number of embryos obtained	Number of β-gal stained embryos	% Chimeric embryos
siMbnl1+2 line 1	45	36	80%
siMbnl1+2 line 2	30	26	87%

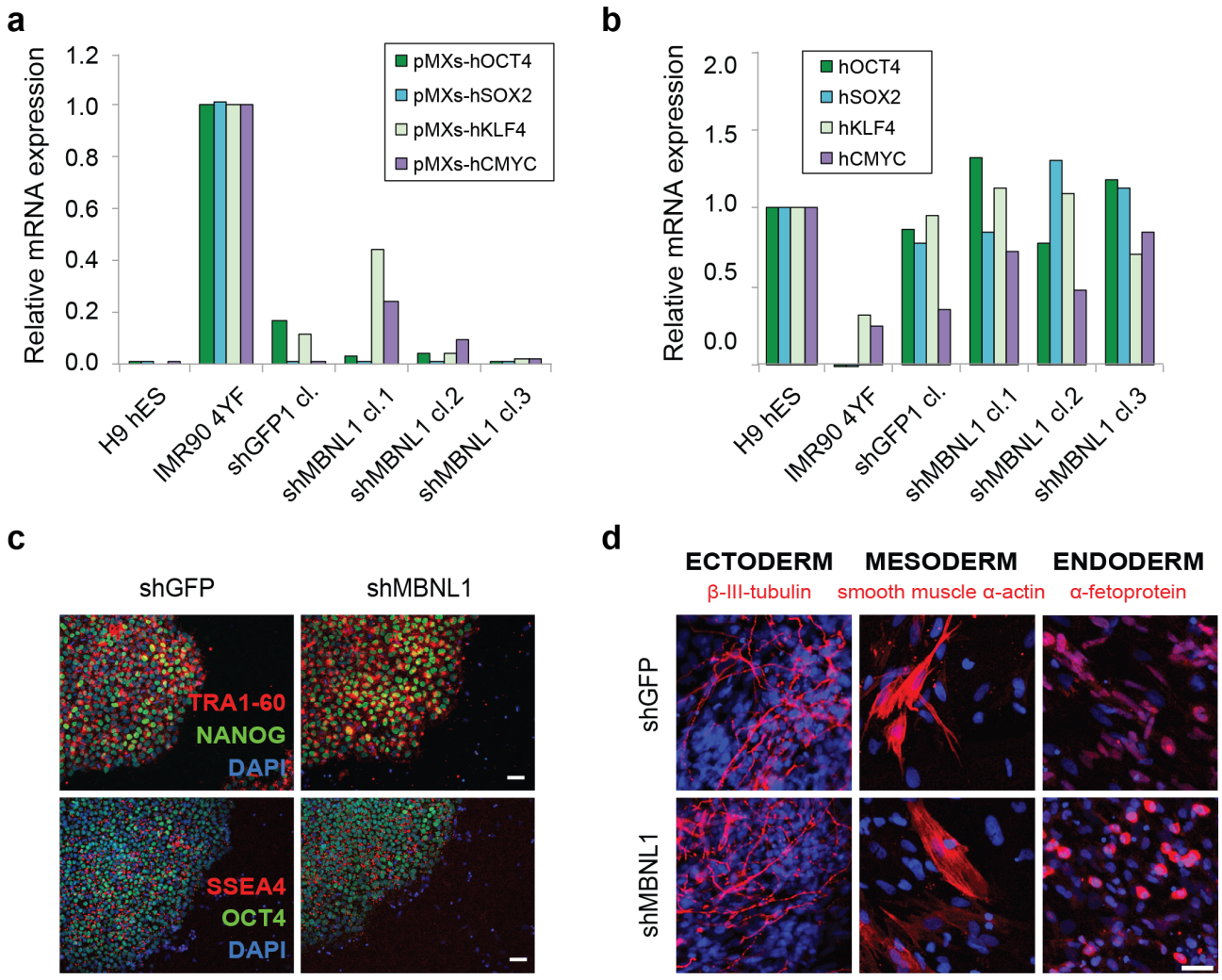
Supplementary Fig. 13 – Chimerism testing for transgene-independent siMbnl1+2 iPSC lines produced from secondary MEF reprogramming

iPSCs derived from secondary MEF reprogramming following siRNA knockdown of Mbnl1+2 contribute to embryonic development. Chimeric embryos were generated by aggregating ICR (CD1 albino) morulae with GFP-positive/LacZ-positive siMbnl1+2 cell lines. Embryos at E10.5 were dissected, fixed and paraffin embedded. Whole-mount GFP detection (a), and LacZ staining (b) of chimeric embryos derived from one of the siMbnl1+2 iPSC lines. c) Information on numbers of chimeric embryos generated from the two independent siMbnl1+2 iPSC lines. d) Transverse sections of the whole-mount stained embryos show contribution of iPSC-derived LacZ-positive cells to derivatives of all three embryonic germ layers. **hg**, hindgut (endoderm); **m**, mesenchyme (mesoderm); **n**, neural tube (ectoderm).



Supplementary Fig. 14 – Analyses of ESC-differential AS events in transgene-independent and transgene-dependent clones from secondary MEF reprogramming

PSI values of ESC-differential alternative exons in fully reprogrammed iPSCs are significantly correlated with those in transgene-independent clones (a) (r= 0.8, p= 3.2 e^{-12}), but not with PSI values of the same events in transgene-dependent clones (b) (r=0.266, p=0.062). PSI values for three transgene-independent and five transgene-dependent clones were averaged over each data point. Only ESC-differential exons with enough coverage in each of the samples and with a ≥ 25 PSI difference between iPSCs and MEFs were included in the analysis. The 95% confidence and prediction intervals are represented by dashed and dotted lines, respectively.



Supplementary Fig. 15 – Characterization of human iPSC lines generated from MBNL1-depleted fibroblasts

a, b) qRT-PCR analyses of iPSC lines that expressed shRNA targeting either GFP (shGFP1) or MBNL1 (shMBNL1) during reprogramming reveal gene expression patterns that are consistent with full reprogramming: compared to control fibroblasts harvested 7 days after retroviral infection (IMR90 4YF), most lines have silenced expression of viral transgenes (a); also, endogenous genes encoded by the reprogramming-factors are expressed at levels similar to those observed in human embryonic stem cells (H9 hES) (b). c) iPSCs generated from fibroblasts expressing shGFP1 have normal colony morphology and express the pluripotency-associated proteins NANOG, TRA1-60, SSEA4, and OCT4 (scale bar = 50 μm). d) Embryoid body assays show that iPSC lines generated from fibroblasts depleted-of MBNL1 differentiate into cells representing all three germ layers (scale bar = 50 μm).